

TESTING ETS

Somewhere in the past, you've walked into a classroom with a pack of your peers on a Saturday morning, carrying sharpened No. 2 pencils in your sweaty palms.

If you've applied to college in the past 50 years, especially ones on the East and West coasts, you've probably taken the SAT. For graduate school, the Graduate Record Examination. For teacher certification, maybe one in the PRAXIS series.

Those rites of passage are changing for many teenagers and young adults today. While high school juniors still take the SAT with pencils, older students now must register to take the GRE or PRAXIS at one of more than 2,500 computerized-testing centers operated by Sylvan Learning Systems or the Educational Testing Service—the creator of gatekeeper exams that have sorted America's educated society for more than five decades.

This year alone, more than 1 million of the almost 13 million exams the ETS administered were on computer—a number projected to grow rapidly, as paper testing likely declines.

"There will be a day when the large majority of our tests will be on computer in some form," says Nancy S. Cole, the president of the ETS.

But getting to that day, so far, has been a trying experience for the 51-year-old nonprofit organization, based on a bucolic campus here just outside Princeton, N.J.

The technological conversion has chipped away at the credibility and financial well-being of the biggest test-maker in the United States. It has reopened the ETS to allegations of monopolistic practices. And it has raised questions about whether its academic culture will survive or be replaced by a new corporate one.

The number of computer-based tests that ETS has given has grown tenfold in the past five years—from about 100,000 in 1994. As of this year, the GRE and the Graduate Management Admission Test are available only on computer. In most of the 165 countries where students take the service's Test of English as a Foreign Language, or TOEFL, the only version is digital. The SAT, the test that has been the financial bedrock throughout ETS' history, remains a pencil-and-paper test and probably will for a long time.

The significant investments needed to create the tests and build testing centers have strained the ETS' \$456 million budget. Unexpected costs—especially in foreign countries—have created a flow of red ink throughout the past decade that Cole says is now about to be halted.

The deficits, in turn, forced the testing service into laying off employees for only the second time in its history, an episode Cole calls "painful" and one that has created ill will.

All of that has led some to suggest that the organization's early forays into computerized testing have been more trouble than they've been worth.

"The computerized tests were rushed into the market to preserve ETS' pre-eminence in testing and to pre-empt competitors—not because they were ready," asserts Robert A. Schaeffer, the public education director of the Center for Fair and

Since venturing into computerized testing, the ETS has faced budget deficits and critics who claim the venerable testing service jumped prematurely into the market.

BY DAVID J. HOFF

Open Testing—or FairTest—a persistent ETS critic that has helped one dissatisfied taker of the computerized GRE earn a refund.

The troubles have put the ETS under the microscope just as it says it's trying to shape its future—and that of all of testing.

Throughout its history, the Educational Testing Service has linked its growth and innovation to computers.

The SAT—originally called the Scholastic Aptitude Test, briefly the Student Assessment Test, and now known officially by its initials—started in 1926 as an exam that Ivy League schools could use to award scholarships to students who didn't hail from elite New England families.

Once it could be given on a mass scale for little cost, the test became a standard hurdle for college admissions. Computerized score sheets allowed the SAT to grow to half a million test-takers by the 1959-60 school year and then to 1 million just five years later. Today, the SAT is given almost 3 million times a year to more than a million students.

Handling tests in such volumes has made the ETS a wealthy—though not-for-profit—institution.

The most visible sign of that wealth is the 376-acre campus in the countryside here outside Princeton. The grounds are part corporate headquarters, part college campus, and part state park.

A long, stately entryway leads to buildings that—in the college tradition—are named for founders and past leaders: Conant Hall (named for founding Chairman James Bryant Conant, a president of Harvard University), the Brigham Library (named for Carl Campbell Brigham, the psychometrician who designed the SAT), and the Chauncey Conference Center (for Henry Chauncey, the founding president and a force behind the ETS' birth in 1948).

As much as the setting resembles a college campus, it also takes after the ETS' corporate neighbor, Bristol, Meyers,

NMENT



Page 2

Use a No. 2 pencil only. Be sure each mark is dark and complete, fills the intended oval. Completely erase any errors or stray marks.

Start with number 1 for each new section. If a section has fewer questions than answer spaces, leave the extra answer spaces blank.

SECTION

3

1	A	B	C	D	E
2	A	B	C	D	E
3	A	B	C	D	E
4	A	B	C	D	E
5	A	B	C	D	E
6	A	B	C	D	E
7	A	B	C	D	E
8	A	B	C	D	E
9	A	B	C	D	E
10	A	B	C	D	E
11	A	B	C	D	E
12	A	B	C	D	E
13	A	B	C	D	E
14	A	B	C	D	E
15	A	B	C	D	E

16 (A) (B) (C) (D) (E)
17 (A) (B) (C) (D) (E)
18 (A) (B) (C) (D) (E)
19 (A) (B) (C) (D) (E)
20 (A) (B) (C) (D) (E)
21 (A) (B) (C) (D) (E)
22 (A) (B) (C) (D) (E)
23 (A) (B) (C) (D) (E)
24 (A) (B) (C) (D) (E)
25 (A) (B) (C) (D) (E)
26 (A) (B) (C) (D) (E)
27 (A) (B) (C) (D) (E)
28 (A) (B) (C) (D) (E)
29 (A) (B) (C) (D) (E)
30 (A) (B) (C) (D) (E)

31	(A)	(B)	(C)	(D)	(E)
32	(A)	(B)	(C)	(D)	(E)
33	(A)	(B)	(C)	(D)	(E)
34	(A)	(B)	(C)	(D)	(E)
35	(A)	(B)	(C)	(D)	(E)
36	(A)	(B)	(C)	(D)	(E)
37	(A)	(B)	(C)	(D)	(E)
38	(A)	(B)	(C)	(D)	(E)
39	(A)	(B)	(C)	(D)	(E)
40	(A)	(B)	(C)	(D)	(E)

If section 3 of your test book contains math questions that are not multiple-choice, continue to item 16 below. Otherwise, continue to item 16 above.

**ONLY ANSWERS ENTERED IN THE OVALS IN EACH GRID AREA WILL BE SCORED.
YOU WILL NOT RECEIVE CREDIT FOR ANYTHING WRITTEN IN THE BOXES ABOVE THE OVALS.**

BE SURE TO ERASE

COMPLETELY.

DO NOT WRITE IN THIS AREA

PLEASE PRINT YOUR INITIALS

$\frac{1}{2} \text{ sec.}$	$\frac{1}{2} \text{ sec.}$	$\frac{1}{2} \text{ sec.}$
----------------------------	----------------------------	----------------------------

More than 1 million of the almost 13 million exams the ETS administered this year alone were via computer—a number expected to grow rapidly.

Squibb, a mile or so away. The headquarters of the drug manufacturer, too, is tucked safely back from the heavily traveled U.S. Route 1.

Throughout the day, hikers, bicyclists, and runners use the nature trails that the testing service opens to the public as part of a state program.

The organization's real estate is valued at \$133.4 million, according to its 1998 tax form filed with the Internal Revenue Service. As a nonprofit, the testing service pays no property tax to Lawrenceville Township. Instead, it donates about \$100,000 a year to a local educational foundation, according to Kevin Gonzalez, an ETS spokesman.

The ETS also held \$34.8 million in cash and \$132 million in stocks and bonds last year, the IRS form says.

While its campus may look like a park or a university, the testing service's executives are paid on the order of business executives. Cole earned \$467,481, plus \$49,665 in deferred compensation, in fiscal 1998. Sharon P. Robinson, the former U.S. Department of Education and National Education Association official who is now the senior vice president and chief operating officer, made \$359,472, \$34,428 in deferred compensation, and had a \$23,200 expense account. Other officers were paid between \$108,097 and \$317,576.

Still, the 2,100-person workforce behaves more like a university faculty than like corporate managers. Its elite corps of researchers are producing cutting-edge research, using the most sophisticated testing techniques and technologies.

ETS researchers regularly publish their work in academic journals and win awards from groups such as the American Educational Research Association and the Center for Research on Education Standards and Student Testing. Its scholars often sit on panels convened by the National Research Council, the respected scientific body that often explores questions in testing and assessment.

"They have some of the best psychometricians in the country," says George F. Madaus, an education professor at Boston College and a leading testing expert.

The testing service's reach in the assessment culture is just as significant as its assets.

In addition to the gatekeeper exams it conducts for the College Board, ETS produces a wide variety of assessments and provides extensive services for schools on every level and professional organizations.

At the same time ETS operates the SAT and PSAT programs that give 5 million tests a year, it also administers the National Board for Professional Teaching Standards' certification exam—to about 4,700 candidates this year. Its stable of exams also includes placement tests given to first-year students in both the University of California and California State University systems.

The ETS processes students' applications

to the University of California, publishes a series of arts education handbooks, and works with the National Urban League to run programs aimed at improving minority achievement.

The testing company also sees itself as providing services that educators need. Early next year, Cole says, ETS plans to release a CD-ROM that will help teachers write classroom tests linked to their states' standards, for example.

Moreover, it owns the Chauncey Group International—a for-profit subsidiary that produces job-certification exams for nurses, architects, and information-technology professionals. Like its parent, Chauncey is dedicated to developing computer-based testing and offering it through Sylvan's testing centers.

Today, if you want to take the GRE, the TOEFL, the Graduate Management Admission Test, or any of the other computerized tests the ETS offers, you skip filling out a registration form and mailing it to the testing service, as was the common practice. Instead, you call the ETS and reserve a seat at a computer in one of almost 2,500 storefronts operated by Sylvan or in a center at more than 100 colleges and universities that the ETS runs on its own.

On the day of the test, you show up at the appointed hour and are assigned a computer, which is programmed to offer a tutorial.

When you start a computer-adaptive test such as the GRE, you'll see questions similar to those on earlier generations of the paper exams. The first will be of moderate difficulty. If you answer correctly, the next will be a little harder.

After you answer each question, the computer adjusts the difficulty of the next question based on how you answered the previous one, gauging your ability at every step along the way. By the end, the computer identifies where you belong on the test scale.

"What this means is that a test-taker doesn't waste time answering questions that are too difficult or too easy," says Linda Cook, the testing service's vice president of assessment.

Cook and other ETS officials say the advantages of computerized testing are many. It takes less time to assess where students fit on the scale. The score is reported to the student immediately. It does not add to the gap in scores between white and minority test-takers, they say, and may even reduce it. And the scores are more reliable than for paper-and-pencil tests.

Critics aren't convinced.

"It's using the computer simply as an elaborate page turner," says Winton H. Manning, a former ETS senior vice president for research and development. The content of the test isn't radically different from what it was on paper, he says, and is still too focused on verbal and mathematical ability. It doesn't do anything to gauge a test-taker's problem-

solving skills beyond those topics, he says.

"There's no reason young people cannot demonstrate their capacity for problem-solving" on the computer, Manning adds.

During his 25-year tenure at the ETS, Manning laid the groundwork for computerized testing. When he retired seven years ago, he feared that such testing "would end up locking the SAT in place" and never lead to performance-based questions that challenge students' critical-thinking skills. In his view, that's exactly what has happened.

Other critics—such as Amy J.C. Cuddy—question whether the computerized tests are even reliable at measuring students' ability.

On Oct. 6, 1998, Cuddy walked into a computer-testing center to take the GRE. She says she had studied "on and off" for six months and "pretty seriously" for the month before. She had practiced using several study guides, including software that the ETS produces and sells. She expected to score in the mid-600s—out of a possible 800—on each of the three sections.

By the end of the day, she was in tears, certain that her hopes of entering a graduate psychology program were ruined. The computer said she had scored 300 on the analytical section, ranking her in the 3rd percentile of test-takers.

Unwilling to let that score stand, Cuddy signed up to take the test with pencil and paper the next month. She scored 690—the 84th percentile—on the analytic section. Her quantitative score rose by another 100 points, from 550 to 650.

As she recalls her experience taking the computerized test, she suspects her early success penalized her. After she answered the first questions correctly, the computer gave her harder questions. As the degree of difficulty increased, so did the time she took to answer.

Once she realized she was running out of time, she started guessing. With a one in five chance of getting the question right, she started to record wrong answers. With every wrong answer, the computer gave her an easier question. By the end of the test, the computer assumed she could not identify the right answers to even the easiest questions. It placed her in the 3rd percentile.

Cuddy, nonetheless, was accepted into the psychology program at the University of Massachusetts at Amherst, where she had been working as a research assistant since her graduation from the University of Colorado at Boulder in 1998. With her application, she included a memo noting she was appealing the computerized score and hoped the admissions counselors would decide based on her written score.

"If I hadn't been in the position I was in



and didn't know people in the field, I probably would have been blackballed by admissions committees," Cuddy says.

After threatening to sue the ETS, Cuddy won a cancellation of her computerized scores and \$357 to refund her testing fee and compensate her efforts in challenging her scores.

Her experience was rare, but it happens enough for ETS researchers to seek ways to avoid it. At a press briefing last month in Washington, Cole announced that the testing service would spend \$6 million to \$8 million to research its computer-adaptive-testing programs. One of the issues, according to Cook, the vice president of assessment, will be to try to help those who are forced to guess at their answers as their time runs out. While Cook didn't mention Cuddy by name, she did describe her experience.

But to critics, such a research initiative only proves that the ETS has been giving tests that aren't necessarily valid or reliable.

"They're describing Amy Cuddy's case," says Schaeffer of FairTest, which helped her secure her refund. "Six years after implementing computerized testing, they're doing research of what the problems are."

"They have such huge sample sizes," Cuddy says. "A few misscored tests are meaningless to them, but they're crushing to the test-takers. The ETS is an unchecked monopoly and students have to deal with it as individuals," she adds.

The monopoly assertion arises often enough that the testing service cites it in the list of frequently asked questions on its World Wide Web site.

"Is ETS an unregulated monopoly?" the ETS Web site asks.

"No on both counts," the site declares. "Although we are the world's largest testing organization, there are competitors for most of the testing programs and related products and services we develop."

While the ETS does have competitors in licensure tests for such fields as teaching and computer-software technicians, it practically has a lock on several testing programs that form the core of its business.

Since the testing service opened in 1948, no other company has held any portion of the SAT contract, and it appears that the ETS will keep its grasp on the college-admissions test over the next several years.

The SAT program has competition from the ACT—the subject-based exam required for admission to most Midwestern and Southern colleges and universities. Despite public perception that the SAT is more popular, about the same number of high school students take each test.

And last year, when the Department of Education requested proposals for a four-year, \$76 million contract to operate the National Assessment of Educational Progress, only the ETS submitted bids for the work in writing test questions, scoring the results, and handling related research.

There haven't been any serious competitors for the contract since the ETS won it away from the Education Commission of the States in 1983, according to Emerson J. Elliot, the U.S. commissioner of educational statistics during much of that era.

"For a new company to bid on what essentially is the same project, you're starting the race from a hundred yards back," says Michael H. Kean, the vice president of public and governmental affairs for CTB/McGraw-Hill, a leading commercial test publisher based in Monterey, Calif. "We could

"We've always known we'd face a period of difficult finances when computer-based-testing volume became large enough to take away from paper volume," ETS President Nancy Cole says.

have done the work, but probably couldn't have been as competitive price-wise."

The ETS' entrance into computerized testing once again has raised suspicions about its power. A federal antitrust trial against Sylvan Learning Systems—the testing service's partner in computerized testing—is scheduled to begin in Cedar Rapids, Iowa, next month.

ACT Inc., the nonprofit company based in Iowa City, Iowa, that runs the ACT admissions tests, alleges that Sylvan undermined its attempt to establish computer-based-testing centers in 1996 by stealing a client away.

ACT officials had been working with the National Association of Securities Dealers for three years to create certification tests. By late 1995, the two organizations were close to concluding an agreement in which ACT would have purchased the association's testing centers and used them to offer tests for other ACT clients, including the American Nurses Credentialing Center.

In early 1996, the lawsuit claims, Sylvan courted the security dealers' association with an anti-competitive bid. In March of that year, according to the suit, representatives from Sylvan and the Chauncey Group International—the three-year-old for-profit ETS subsidiary—met with an ACT client. They purportedly said that the Iowa nonprofit

The not-for-profit corporation's operations—and bucolic setting— have much in common with the academic culture.

would be unable to satisfy the client's needs for computer-based testing. Moreover, according to the suit, Sylvan representatives said their company would only allow a computer-based test developed by Chauncey to be offered in its centers.

While neither ETS nor Chauncey is a party to the lawsuit, Cole says that she and several other testing-service officials have given depositions.

Beyond the questions of how the ETS conducts its business, the testing service's entrance into computer-based testing raises the issue of whether such testing has been good business.

In the fiscal year that ended June 30, the ETS had a \$7.7 million operating deficit. (Revenues used to calculate the deficit exclude interest or dividends generated from the \$165 million in cash and stock accounts that the ETS has among its assets, Cole says.) The year before, the ETS ran a deficit of \$8 million. Since 1990, its budget has been balanced once and shown a surplus once.

"We've always known we'd face a period of difficult finances when computer-based-testing volume became large enough to take away from paper volume, so we're running two delivery systems in parallel without the numbers to support either of them," Cole says. "When it hit, we found we weren't as prepared as we hoped we would be."

As paper-based testing declined, the ETS lost the economies of scale it offers, forcing it to seek cuts elsewhere.

Every section of the organization was asked to find ways to cut back.

The assessment division created new software that it says reduces the cost of creating test questions, according to Cook. Chief Financial Officer Frank Gatti instituted purchasing policies to get better deals from suppliers. The division in charge of the SAT has met \$17 million in cost-cutting targets called for in its current three-year contract with the College Board.

The hardest of the cutbacks was to let go 60 employees and leave another 60 or so positions unfilled.

The actions upset the ETS culture, where getting a job was akin to earning tenure at a college. The only other major layoffs in the nonprofit's history were in 1983, when it lost a contract to process federal student-loan applications.

"ETS used to be a job for life," says Kean, the CTB/McGraw-Hill vice president and a former manager of the Midwest office of the ETS. "They were good people, mind you. But it was very much like having tenure at a university."

While trimming elsewhere, the costs of computer-based testing mushroomed out of control.



On a computerized version of the GRE, Amy J.C. Cuddy scored 300 on the analytic section. The next month, taking a paper version, she scored 690.

The biggest problems happened overseas, where the ETS had planned to eliminate the paper testing of the TOEFL, the test international students take to demonstrate their English proficiency to U.S. universities.

In working with Sylvan to create testing centers in developing countries, the ETS watched costs rise sharply. Because much of the Third World experiences daily electrical breakdowns, testing centers had to purchase backup generators to keep computers running.

Where permanent centers were not yet in place, the groups created temporary ones, often in airports.

Despite the money sunk into the project, the ETS couldn't reach its goal of offering all TOEFL exams on computer. The tests continue to be given on paper in some regions in Africa.

Cole maintains that the testing service's finances are improving. The \$7.7 million deficit in the past fiscal year was significantly lower than the \$18 million projected at the start of the year, she says.

In a September newsletter for ETS employees, CFO Gatti reported that the organization had "a good chance" of breaking even in fiscal 2000—a year ahead of schedule.

"They have improved the situation this year even better than they set out to do," says Henry Chauncey, the ETS' president for its first 22 years, who now, at age 94 is retired in Shelburne, Vt. "They have a reserve. They're in good shape."

In large measure, the ETS can look to its stalwart as its savior.

The testing service's financial situation looked shakier earlier this year as rumors spread that the new leadership at the College Board was considering subcontract-

ing portions of the SAT.

Those rumblings have subsided for now. Cole and Gaston Caperton, who became the president of the New York City-based College Board this past summer, say the two sides are close to agreeing to a new contract that would keep the ETS as the sole contractor for the near future.

"It is a unique relationship," says Caperton, a former governor of West Virginia and insurance executive. "I've never seen a relationship like it in my experience in business and government."

Caperton says he has focused on improving the way the nonprofits work together, not ending their relationship.

That relationship has existed since the ETS was founded. The testing activities of the College Board, the American Council on Education, and the Carnegie Corporation of New York were all combined under the ETS umbrella, according to Mr. Chauncey, who helped forge the compromise.

Since then, the ETS has managed all the work of the SAT, from designing the test questions to printing the test booklets to reporting the scores.

Contracts for the SAT and other College Board programs, such as the Advanced Placement exams, amounted to almost \$160

million last year, according to the ETS annual report.

"The relationship between the College Board and ETS is so much centered on the SAT, with the two organizations having in effect joint ownership of the test, that it would be impossible to break that relationship apart," says Donald M. Stewart, a former College Board president who now is a senior program officer for the Carnegie Corporation. "I can't imagine the two organizations coming apart in any way or changing the relationship. To walk away from it would be silly."

Cole, however, does acknowledge that the relationship may evolve. "I suspect that 50 years from now, we'll still be working with the College Board," she says. "We have such common history and common goals. But there will be aspects of it that will change."

If that's the case, it may bode well for the financial future of the ETS, some experts say. The SAT is an inexpensive test to administer, creating excess revenues to be directed elsewhere in the organization.

"The hard fact is that ETS is dependent financially on the SAT, and so is the College Board," says Manning, the former ETS research chief. "It throws off enough profit so they can do a bunch of other things that are worthwhile doing."

Still, the SAT doesn't necessarily fit into the ETS' long-term strategy of computerizing tests. The exam has such a high volume that it would be too costly to purchase the hardware and draft the questions and still keep the price affordable, Caperton and Cole say.

"I don't see how [computerized testing in the SAT] is to anyone's advantage," Caperton says. "I don't see how it reduces costs of administering the tests."

While the Educational Testing Service may have made some missteps in the entry into computerized testing, many say it's headed in the right direction.

"They put out a good product," says Madaus of Boston College. "It's fallible, but all tests are fallible. I think they would recognize that."

"Computerized testing is inevitable," Stewart says. "ETS may have rushed into it more quickly than they should have, but they're on the right track."

Where that path will lead is up in the air. Even the organization's top researchers don't know exactly what testing will look like in 2048, when the ETS would celebrate its 100th anniversary.

In a 1998 paper titled "Reinventing Assessment," ETS researcher Randy Elliot Bennett predicts there will be three stages of computerized testing.

In the first stage, which the ETS and others are deep into, he says, the new tests are on computer, but "they are substantively the same as those administered on paper: They measure the same skills, use the same behavioral designs, and de-

pend primarily on the same types of tasks," Bennett writes.

In the second phase, the tests will be "qualitatively" different, in Bennett's scenario. For example, a multimedia test might ask an Advanced Placement history student to analyze and explain the significance of clips from World War II propaganda films.

The costs of computerized testing may become reasonable because the computers themselves might be able to write test questions as students take the test.

In the final stage, students will no longer sit for specific exams, but will be tested repeatedly as they learn material from a prescribed curriculum.

"Dedicated test centers also may be on the endangered list," he writes. The need for large-scale tests—such as the SAT—may de-

cline because much of its content might be embedded into tests tied to curriculum.

There will be the potential that wireless technologies will make it easy to cheat: It might become possible to intercept the questions and communicate the correct answers to a test-taker sitting in front of a computer screen.

Bennett doesn't offer a time line for the three stages.

For the leadership of the testing service, the longer the better. For all the struggles ETS researchers and officials had in creating the current generation of computerized tests, they don't know how they will be able to manage life in Bennett's final stage.

"We haven't a clue what business model will allow it to support itself," Cole says.

And for now, at least, the ETS has no multiple-choice list from which to choose. ■

By the Number: ETS Assessments

Below are the tests that the ETS administers, including those for the College Board, its biggest client. The two have been associated since the ETS was founded in 1948.

	Projected Volume 1999-2000
COLLEGE BOARD	
• SAT	2.9 million
• Preliminary SAT	2.1 million
• Accuplacer	1.4 million
• Advanced Placement	1.2 million
• College-Level Examination Program	120,300
• Multiple Assessment Programs and Services	117,100
• Academic Profile	74,400
• Algebridge	13,200
OTHER TESTS	
• PRAXIS: Professional Assessments for Beginning Teachers	813,483
• Test of English as a Foreign Language	689,500
• Graduate Record Examinations	474,000
• National Assessment of Educational Progress	470,000
• Education Records Bureau	261,600
• Graduate Management Admission Test	201,000
• California State University: English & Mathematics Placement Tests	84,000
• GRE: Major Field Tests	51,400
• Test of Spoken English	22,800
• PLUS Academic Abilities Assessment	21,500
• University of California Writing Exam	17,900
• Secondary Level Proficiency Test	9,200
• High Schools That Work	8,000
• Prueba de Admision para Estudios Graduados	6,700
• National Board for Professional Teaching Standards	4,700
• School Leaders Licensure Assessment	1,330
• Tasks in Critical Thinking	1,140

SOURCE: Educational Testing Service